Chapter 8.

Continuous Fields, Time Dependence, Infinite 3D

Michigan Avenue
Robert Wittig

Contents

8.1. Introduction

One might, with good reason, assume that only modest effort is needed to go from the lattice of discrete mass points introduced in Fig. 1 of Chapter 1 to the corresponding continuous mass field. Figure 1 illustrates how the progression toward yet finer mass grids takes us to the desired limit. Despite the conceptually straightforward nature of the switch from discrete masses to a continuous mass field, a patient approach will prove its worth.

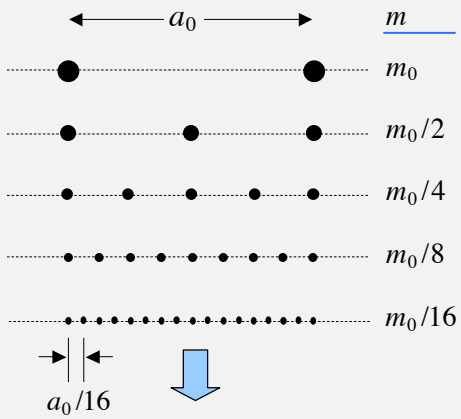


Figure 1. The ratio m/a retains its value as m and a each approach zero.

This is an especially important section because displacement and momentum fields make their entry. The term *field* in the present context means something that varies continuously throughout space. For example, a static electric field has magnitude and direction that vary continuously in 3D, as illustrated in Fig. 2. Likewise, a second-rank tensor field that describes dielectric polarization varies continuously in 3D on length scales that exceed greatly those of the unit cells. If time variation is included, the fields vary continuously in 4D spacetime.

Continuous fields will be encountered in the present chapter, Chapter 9, and Part V. These fields are continuous by their nature. That is, they have nothing to do with a discrete lattice or any approximation. The electromagnetic field, as well as massive fields such as Klein-Gordon and Dirac will be quantized – the electromagnetic and Klein-Gordon fields using commutators, as was done with phonons, and the Dirac (fermion) field using anticommutators, for reasons that will be explained later.

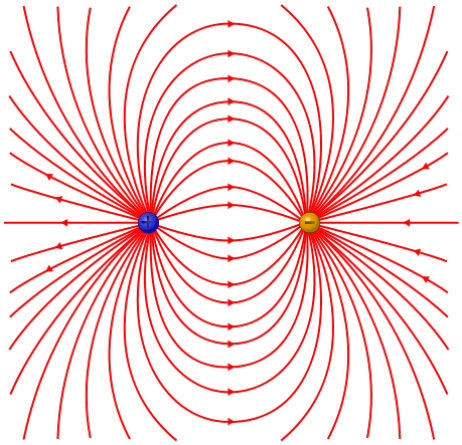


Figure 2. Cross sectional view: lines of constant electric field strength are indicated for + Q (left) and – Q (right).

As mentioned earlier, the transition from the discrete lattice to a continuous mass distribution requires that the mass density m/a retain its value as $a \to 0$ and $m \to 0$, as indicated in Fig. 1. In addition, the constant: $v = \Omega a$, which is the speed of sound in the small wave vector region of the first Brillouin zone, as well as in the limit of a continuous mass distribution (in which case, the edges of the first Brillouin ($\pm \pi/a$) zone go to $\pm \infty$), also retains its value.

What is not obvious *a priori* is how the displacement and momentum fields scale with discrete mass m and lattice spacing a. In other words, in making the transition from the discrete lattice operators $\hat{q}_n$ and $\hat{p}_n$ (which operate on the mass point that is harmonically bound to the location $x = na$ of a rigid lattice; see Fig. 1 in Chapter 1) to the continuous density operators $\hat{q}(x)$ and $\hat{p}(x)$, how are m and a subsumed into these operators?

In mathematical terms, the relations: $\hat{q}(x) = f(m,a)\hat{q}_n$ and $\hat{p}(x) = g(m,a)\hat{p}_n$ are sought. The densities $\hat{q}(x)$ and $\hat{p}(x)$ must subsume the mass, each in its own way, such that mass no longer appears explicitly in the Hamiltonian. Likewise, the original lattice spacing a must also be subsumed. It cannot appear in the final expressions because with a continuous mass distribution it no longer exists. Because $\hat{q}_n$ and $\hat{p}_n$ appear quadratically in the Hamiltonian, it is anticipated that the factors $f(m,a)$ and $g(m,a)$ will be square roots of some kind.

To determine these factors, the Hamiltonian is examined in the limits: $a \to 0$ and $N \to \infty$, with the restriction that the speed of sound, $v = \Omega a$, and the mass density, m/a, are each held fixed. In addition, the product Na must be equal to the length of the lattice, L, and locations na must give way to the continuous parameter x. Hereafter, this will be referred to as the *elastic limit*.

Referring to the Hamiltonian given by eqn (1.3) in Chapter 1, another important point is that the respective amounts of kinetic and potential energy should be unchanged as the elastic limit is taken. In other words, when summation of $\hat{p}_n{}^2/2m$ over n is replaced by integration from 0 to L of a kinetic energy density, integration must yield the same value as summation … and likewise for the potential energy.

8.2. Continuous Mass Distribution

The expression for the Hamiltonian given by eqn (1.3) in Chapter 1 is now rewritten using $\Omega^2 = v^2/a^2$, and with a factor of a multiplying the kinetic and potential energy terms in anticipation of summation being converted to integration:

$$\hat{H} = \tfrac{1}{2}\sum_{n=1}^{N}\left(\frac{\hat{p}_n{}^2}{m} + m\Omega_0{}^2\,\hat{q}_n{}^2 + m\Omega^2\left(\hat{q}_n - \hat{q}_{n-1}\right)^2\right)$$

$$\Rightarrow \quad \tfrac{1}{2}\sum_{n=1}^{N} a\left(\frac{\hat{p}_n{}^2}{ma} + \Omega_0{}^2\,\frac{m}{a}\hat{q}_n{}^2 + v^2\,\frac{m}{a}\left(\frac{\hat{q}_n - \hat{q}_{n-1}}{a}\right)^2\right). \tag{8.1}$$

In taking the elastic limit of the second line in eqn (8.1): (i) the term $[(\hat{q}_n - \hat{q}_{n-1})/a]^2$ becomes $(\partial\hat{q}_n/\partial x)^2$; (ii) the second and third terms within the large parentheses retain their values in the elastic limit because m/a and v^2 are each constant, and $\hat{q}_n$ is allowed to retain its form for the time being; and (iii) the first term retains its value because $\hat{p}_n$ is proportional to mass, and consequently $\hat{p}_n{}^2/ma$ is proportional to m/a, which is constant. As with the discrete displacement operator $\hat{q}_n$, the discrete momentum operator $\hat{p}_n$ is allowed to retain is form for the time being. Let us now introduce the elastic limit by replacing summation with integration

$$\lim_{a\to 0,\, N\to\infty} \left\{ \sum_{n=1}^{N} a(\cdots) \right\} = \int_0^L dx(\cdots) , \tag{8.2}$$

where $na \to x$, $a \to dx$, and $Na = L$ have been used. Putting all of the above stuff into eqn (8.1) yields

$$\hat{H} = \tfrac{1}{2} \int_0^L dx \left(\frac{\hat{p}_n{}^2}{ma} + \Omega_0{}^2 \frac{m}{a} \hat{q}_n{}^2 + v^2 \frac{m}{a} \left(\frac{\partial \hat{q}_n}{\partial x} \right)^2 \right). \tag{8.3}$$

The Hamiltonian is now expressed as the integral of a *Hamiltonian density* that is equal to one-half times the contents of the large parentheses. Equation (8.3) appears to be a hybrid because the operators $\hat{p}_n$ and $\hat{q}_n$ are not densities. It remains to replace $\hat{p}_n$ and $\hat{q}_n$ by their density counterparts, and in so doing eliminate the explicit appearance of m and a. From eqn (8.3) it is clear how the aforementioned scaling works. Namely, the expressions for the momentum and displacement fields are

$$\hat{p}(x) = (ma)^{-1/2} \hat{p}_n \qquad \hat{q}(x) = (m/a)^{1/2} \hat{q}_n . \tag{8.4}$$

The constancy of the displacement field as the elastic limit is taken is illustrated in Fig. 3. Note that the magnitude of the displacement (blue arrows) does not change in going from top to bottom. All of the relationships needed to convert the particle Hamiltonian of eqn (8.1) to one involving the Hamiltonian density are now in place. The Hamiltonian density is represented in terms of $\hat{q}(x)$ and $\hat{p}(x)$, and neither the discrete mass nor the lattice spacing appears in the expression for the Hamiltonian. These parameters no longer exist as discrete units. They are distributed according to the densities $\hat{q}(x)$ and $\hat{p}(x)$. The Hamiltonian given by eqn (8.3) now reads

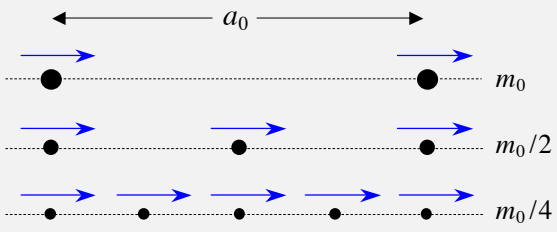


Figure 3. In the elastic limit, displacement stays the same as m and a approach zero with m/a fixed. It is distributed over a continuous mass field.

$$\hat{H} = \tfrac{1}{2} \int_0^L dx \left(\hat{p}(x)^2 + \Omega_0{}^2 \hat{q}(x)^2 + v^2 \left(\frac{\partial \hat{q}(x)}{\partial x} \right)^2 \right). \tag{8.5}$$

Next, the commutator relation: $[\hat{q}_n, \hat{p}_{n'}] = i\delta_{n,n'}$ ($\hbar = 1$), is converted to the analogous one for the operators $\hat{q}(x)$ and $\hat{p}(x')$. In so doing, we must take care with the right hand

side of the commutator relation. With $[\hat{q}_n, \hat{p}_{n'}]$, the right hand side is $i\delta_{n,n'}$, whereas with $[\hat{q}(x), \hat{p}(x')]$, the right hand side must ensure that $x = x'$. In other words, it is necessary to satisfy the requirement that physics is local. At equal times an event at x_1 cannot affect a particle located at x_2. In relativistic language, x_1 and x_2 are related by a space-like interval. As a result, there can be no causal relationship between them. Likewise, at a given location, an event at t_1 has immediate consequence, not at a later or earlier time t_2. This equal time requirement also follows from special relativity, as space and time must be treated on equal footing. You may wish to ponder this.

Mathematically, ensuring that physics is local is the role of the Kronecker delta $\delta_{n,n'}$ and the delta function $\delta(x - x')$. The distinction between these two deltas is that $\delta_{n,n'}$ deals with discrete quantities such as mass points and their momenta, whereas $\delta(x - x')$ deals with continuous distributions (fields) such as $\hat{q}(x)$ and $\hat{p}(x)$. You have familiarity with the use of the Kronecker delta with orthogonal functions. These are used with a discrete representation on a space. The delta function $\delta(x - x')$ serves the same purpose, though the representation is continuous.

The discrete mass points obey the relation $[\hat{q}_n, \hat{p}_{n'}] = i\delta_{n,n'}$, and this cannot depend on the distances between the mass points, that is, the size of a. Again, this is a statement of local physics. The mass point at n obeys $[\hat{q}_n, \hat{p}_n] = i$, whereas the one at $n+1$ obeys $[\hat{q}_{n+1}, \hat{p}_{n+1}] = i$. Each particle obeys its own commutation relation – not to be shared with another particle. In the elastic limit, the fields $\hat{q}(x)$ and $\hat{p}(x)$ must retain this local character. This might seem curious at first sight because quantum mechanics prevents particles from being localized. They obey $\Delta x \Delta p \geq 1/2$, so how can we demand that x and x' are the same, which is exactly what $\delta(x - x')$ does?

What is happening is that x is a *parameter* in the field operator. It has been demoted from its role of observable in the quantum mechanics of particles, where 1D location is indicated with the operator $\hat{x}$. In this sense it has been placed on equal footing with time. Neither are observables. The observables are represented with $\hat{q}$ and $\hat{p}$. The value of, say, $\hat{q}$ depends on where it is. In other words, it depends on x, so we write $\hat{q}(x)$. This is like the displacements in Fig. 8.3, which are present on a progressively finer mass grid as the elastic limit is taken. The displacements $\hat{q}(x)$ constitute a field on x.

The mathematical realization of the above statements is that the Kronecker delta $\delta_{n,n'}$ is replaced by the delta function $\delta(x - x')$, which is a density. The transformation from $\delta_{n,n'}$ to $\delta(x - x')$ works as follows (Fig. 4):

$$[\hat{q}(x), \hat{p}(x')] = \lim_{a \to 0,\, N \to \infty} \left[\sqrt{\frac{m}{a}} \hat{q}_n,\ \frac{1}{\sqrt{ma}} \hat{p}_{n'} \right]. \tag{8.6}$$

The m's cancel, and using $[\hat{q}_n, \hat{p}_{n'}] = i$ yields

$$[\hat{q}(x), \hat{p}(x')] = i \left(\lim_{a \to 0,\, N \to \infty} \frac{\delta_{n,n'}}{a} \right) = i\,\delta(x - x'). \tag{8.7}$$

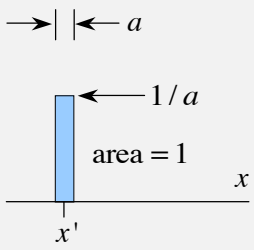


Figure 4. The delta function is obtained from the Kronecker delta through a limiting process. The ratio $\delta_{n,n'}/a$ is identified as a density centered at $x' = na$.

$\frac{\delta_{n,n'}}{a} \to \infty$ as $a \to 0$ in a way such that the product $\left(\frac{\delta_{n,n'}}{a}\right) a = 1$ as $a \to 0$.

$$\int_{x'-\varepsilon}^{x'+\varepsilon} dx\, \delta(x - x') = 1$$

where $\delta(x - x') = 0$ for $|x'| > \varepsilon$

The limit is the delta function $\delta(x - x')$. It is also referred to as the Dirac delta or the Dirac delta function.[1] As mentioned above, it handles functions that have continuous representations on a space in a manner that is analogous to how the Kronecker delta handles functions that have discrete representations on a space.

"In quantum theory whenever an improper function appears, it will be something that is to be used ultimately in an integrand."
Paul Dirac

Great discussions are in Dirac's book: *The Principles of Quantum Mechanics* [68], and Kusse and Westwig: *Mathematical Physics for Engineers and Scientists* [36]. The most thorough discussion (yet accessible to non-mathematicians) I am aware of on the topic of generalized (also called improper) functions, which includes the delta function, is in a book by Lighthill [69]. It has been out of print for some time, but I have a photocopy if you are interested.

Thus, we have obtained an important result. Moreover, this has been achieved without having to work through a great deal of math:

$$\left[\hat{q}(x), \hat{p}(x')\right] = i\delta(x - x'). \tag{8.8}$$

It goes without saying that $[\hat{q}(x), \hat{q}(x')] = 0 = [\hat{p}(x), \hat{p}(x')]$.

---

[1] The delta function under consideration here was introduced well before the birth of Paul Dirac (1902-1984). The eccentric, self-taught English electrical engineer and mathematician Oliver Heaviside (1850-1925) used it in the nineteenth century, and the Dutch mathematician Thomas Stieltjes (1856-1894) was an early proponent. The Slovenian mathematician Josip Plemelj derived what is now frequently called *Dirac's formula*. Dirac made frequent use of the generalized function that came to be called the delta function, and his name has become associated with it. Perhaps English speakers found it hard to pronounce Plemelj or Stieltjes.

Next, the Hamiltonian is transformed to its k-space representation. With the discrete lattice advantage was gained through the k-space (Q_k / P_k) representation. It reduced the problem from one in which it would have been necessary to deal with N nearest neighbor interactions and N site interactions to the simpler problem of pairs of k values.

We shall now carry out the analogous transformation from the coordinate representation to the k-space representation to gain like advantage. As in the case of the discrete lattice, you will see that the results are obtained with no significant mathematical overhead. The reason everything falls into place so nicely is that k remains discrete. This is an important point, but a subtle one. You might wonder about k being discrete. After all, we just finished showing that the number of k values goes to infinity for a continuous mass distribution. That is, as a approaches zero the edge of the first Brillouin zone moves out to $k = \infty$. The discrete nature of the lattice has been removed. Does this not mean that, in the $a \to 0$ limit, k is becoming a continuously variable parameter?

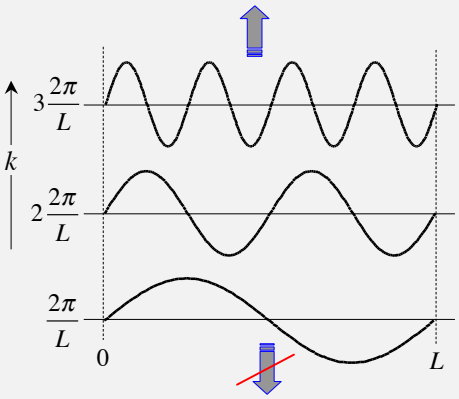


Figure 5. k can increase indefinitely by integer multiples of $2\pi/L$. However, it cannot decrease below $2\pi/L$, with the exception of $k = 0$. These larger wavelengths are not compatible with the boundary conditions.

The answer is no. Though k is allowed to approach infinity as $a \to 0$, it cannot get arbitrarily small because the lattice is of finite length. In other words, one wavelength can match a length L, but wavelengths larger than L cannot be used. As illustrated in Fig. 5, they simply do not fit into the interval $0 - L$, while at the same time satisfying the periodic boundary condition that their corresponding complex wave amplitudes have the same value at 0 and L.[2] The distribution of k values might appear smooth when k is large, but in fact it is unchanged from its discrete lattice counterpart except that many k values are allowed. After all, k can approach infinity. In other words, the k values are still given by $(2\pi / L)l$, where l is an integer.

Another way of saying this is that the k-space remains discrete. Though x varies continuously, k is still given by integer multiples of the fundamental unit $2\pi / L$. As mentioned above, this enables us to obtain the results for the continuous mass distribution in a few lines. For example, the Fourier expansion of $\hat{q}(x)$ in terms of annihilation and creation operators is obtained using results from Chapter 1. Namely, the expansion of $\hat{q}_n$ in terms of $\hat{Q}_k$, and $\hat{Q}_k$, in turn expressed in terms of $\hat{a}_k$ and $\hat{a}^\dagger_{-k}$, is given by

[2] The exception is $\lambda = \infty$. In other words, $k = 0$ satisfies the periodic boundary condition because it has constant magnitude and no phase variation. This mathematical point does not have consequence in the present context.

$$\hat{q}(x) = \sqrt{\frac{m}{a}}\,\hat{q}_n = \sqrt{\frac{m}{a}}\left(\underbrace{\frac{1}{\sqrt{mN}}}_{L^{-1/2}}\sum_k \underbrace{e^{ikan}}_{e^{ikx}}\hat{Q}_k\right) \qquad \frac{1}{\sqrt{2\omega_k}}\left(\hat{a}_k + \hat{a}_{-k}^{\dagger}\right)$$

$$\Rightarrow \quad \hat{q}(x) = L^{-1/2}\sum_k e^{ikx}\frac{1}{\sqrt{2\omega_k}}\left(\hat{a}_k + \hat{a}_{-k}^{\dagger}\right) \tag{8.9}$$

The locations given by na in the discrete lattice model have been replaced by x, which varies continuously. The analogous expression for $\hat{p}(x)$ follows in a like manner. This manipulation and verification of some of the steps discussed below are left as exercises.

To evaluate the commutator $[\hat{Q}_k, \hat{P}_{k'}]$ in the elastic limit, recall that, for the discrete lattice, the relation $[\hat{Q}_k, \hat{P}_{k'}] = i\delta_{k,k'}$ follows from $[\hat{q}_n, \hat{p}_{n'}] = i\delta_{n,n'}$. Most importantly, it was found that this result does not depend on how many n values are involved. No matter how large N becomes, the relation $[\hat{Q}_k, \hat{P}_{k'}] = i\delta_{k,k'}$ still holds. That is, k remains discrete.

Consequently, in the elastic limit, the k-space commutation relations are unchanged: $[\hat{Q}_k, \hat{P}_{k'}] = i\delta_{k,k'}$ and $[\hat{Q}_k, \hat{P}_{k'}] = 0 = [\hat{P}_k, \hat{P}_{k'}]$. Likewise, the k-space representation of the Hamiltonian is unchanged, so its form is identical to the one given by eqn (1.42) in Chapter 1: $\hat{H} = \sum \omega_k(\hat{a}_k{}^{\dagger}\hat{a}_k + 1/2)$, where $\hat{a}_k{}^{\dagger}\hat{a}_k$ is the number operator that acts in Fock space and returns the number of quanta.

The dispersion relation is equally straightforward. Specifically, start with

$$\omega_k{}^2 = \Omega_0{}^2 + 4\Omega^2 \sin^2\left(\tfrac{1}{2}ka\right)$$

$$= \Omega_0{}^2 + (\Omega a)^2\left(\frac{\sin\left(\frac{1}{2}ka\right)}{\frac{1}{2}a}\right)^2.$$

Now take the limit $a \to 0$ and use $v = \Omega a$. This yields $\omega_k{}^2 = \Omega_0{}^2 + v^2k^2$. The above facts are summarized below.

$$\hat{H} = \sum_k \omega_k\left(\hat{a}_k{}^{\dagger}\hat{a}_k + \tfrac{1}{2}\right) \text{ (positive } k\text{)} \tag{8.10}$$

$$\omega_k{}^2 = \Omega_0{}^2 + v^2k^2 \tag{8.11}$$

$$\left[\hat{Q}_k, \hat{P}_{k'}\right] = i\delta_{k,k'} \quad \left[\hat{Q}_k, \hat{Q}_{k'}\right] = 0 = \left[\hat{P}_k, \hat{P}_{k'}\right] \tag{8.12}$$

Keeping in mind that k is discrete, the field operators $\hat{q}(x)$ an $\hat{p}(x)$ are given by

$$\hat{q}(x) = L^{-1/2} \sum_k e^{ikx} \left(\frac{1}{\sqrt{2\omega_k}} \right) \left(\hat{a}_k + \hat{a}_{-k}^\dagger \right) \tag{8.13}$$

$$\hat{p}(x) = L^{-1/2} \sum_k e^{-ikx} \left(-i \sqrt{\frac{\omega_k}{2}} \right) \left(\hat{a}_{-k} - \hat{a}_k{}^\dagger \right). \tag{8.14}$$

Before moving on to time dependence, let us tidy the nomenclature. Referring to eqn (8.13), the sum over k involves $e^{ikx} \hat{a}_{-k}^\dagger$. However, positive and negative k values are distributed symmetrically about zero in the Brillouin zone. Therefore the sum over k of $e^{ikx} \hat{a}_{-k}^\dagger$ is equal to the sum over k of $e^{-ikx} \hat{a}_k{}^\dagger$. Of course, the same idea applies to $\hat{p}(x)$ given by eqn (8.14). Thus, eqns (8.13) and (8.14) become

$$\hat{q}(x) = L^{-1/2} \sum_k \frac{1}{\sqrt{2\omega_k}} \left(e^{ikx} \hat{a}_k + e^{-ikx} \hat{a}_k{}^\dagger \right) \text{ (positive } k) \tag{8.15}$$

$$\hat{p}(x) = L^{-1/2} \sum_k -i \sqrt{\frac{\omega_k}{2}} \left(e^{ikx} \hat{a}_k - e^{-ikx} \hat{a}_k{}^\dagger \right) \text{ (positive } k). \tag{8.16}$$

Referring to eqns (8.10) – (8.12), the fact that the Hamiltonian and commutation relations are unchanged in this continuum limit comes as no surprise. The dispersion relation is more interesting. For example, with $\Omega_0 = 0$ we have $\omega_k = vk$, albeit with discrete k. This describes waves traveling unimpeded in a homogeneous medium. The waves undergo no change of shape as they propagate. This aspect is analogous to an electromagnetic wave in free space, though with speed v instead of c.

8.3. Time Dependence

We shall now examine the time dependence of the displacement and its corresponding momentum, starting with the classical discrete lattice, advancing to the continuous mass distribution, and finishing with the quantum mechanical version. This completes the description of $\hat{q}(x,t)$ and $\hat{p}(x,t)$ for the 1D, finite-length continuous mass distribution. The next steps are: (i) go from 1D to 3D; (ii) allow the volume go to infinity; and (iii) extend the formalism to include complex scalar fields. It is perhaps not obvious at the moment because of the factors e^{ikx} and e^{-ikx}, but the displacement and momentum fields $\hat{q}(x)$ and $\hat{p}(x)$ are real.[3]

[3] To give a hint of why this is so, go back to eqn (8.15). It is obvious that $\hat{q}(x) = \hat{q}(x)^\dagger$, as the parenthetic term is $e^{ikx} \hat{a}_k + (e^{ikx} \hat{a}_k)^\dagger$. That is, the field operator is Hermitian. In classical language we would say $q(x) = q(x)^*$.

Referring to Fig. 1 and eqn (1.1) in Chapter 1, the equations of motion that describe the classical dynamics of the discrete lattice mass points are

$$m\ddot{q}_n = -\kappa_0 q_n - \kappa\big((q_n - q_{n-1}) + (q_n - q_{n+1})\big), \tag{8.17}$$

or, dividing by m, and using $\Omega_0{}^2 = \kappa_0 / m$ and $\Omega^2 = \kappa / m$:

$$0 = \ddot{q}_n + \Omega^2\big((q_n - q_{n-1}) + (q_n - q_{n+1})\big) + \Omega_0{}^2 q_n . \tag{8.18}$$

In the limit of a continuous mass distribution, this becomes

$$0 = \ddot{q}(x,t) - v^2\left(\lim_{a\to 0}\frac{1}{a}\left(\left(\frac{q_{n+1} - q_n}{a}\right) - \left(\frac{q_n - q_{n-1}}{a}\right)\right)\right) + \Omega_0{}^2 q(x,t), \tag{8.19}$$

where $\Omega^2 a^2 = v^2$ has been used. The limit inside the largest parentheses yields a second order partial derivative with respect to x. Thus, eqn (8.19) becomes

$$0 = \frac{\partial^2 q(x,t)}{\partial t^2} - v^2\frac{\partial^2 q(x,t)}{\partial x^2} + \Omega_0{}^2 q(x,t). \tag{8.20}$$

This is a wave equation is easy to solve. With $\Omega_0 = 0$, it is the homogeneous equation of an elastic rod. The first two terms, taken together and divided by v^2, is referred to as the d'Alembertian, named after the French mathematician Jean le Rond d'Alembert (1717-1783). This operator arises so frequently that it has acquired its own symbol, the square $\Box$. Think of the four sides of the square symbolizing the space-time coordinates (ct, x^1, x^2, x^3), just as the three sides of the Laplacian (denoted Δ or ∇^2) symbolize the space coordinates of this operator. It also has some special properties that we will get to in Chapter 9 and throughout Part V.


Jean le Rond d'Alembert

The solution of eqn (8.20) can be expressed as a linear combination of complex exponentials:

$$c_1 e^{i(kx - \omega_k t)} + c_2 e^{-i(kx - \omega_k t)} + c_3 e^{i(kx + \omega_k t)} + c_4 e^{-i(kx + \omega_k t)}, \tag{8.21}$$

where $\omega_k{}^2 = \Omega_0{}^2 + v^2 k^2$. The four coefficients c_i are determined by four conditions, for example, two for the value of the amplitude at two spatial boundaries and at a specific time (say $t = 0$), and two for the time derivative of the amplitude, which is usually taken at the same time and spatial locations. Too many conditions need to be satisfied with the classical solution to have consistency with quantum mechanics. In quantum mechanics, a particle's location and momentum cannot be known simultaneously.

Classically, the time derivative of the displacement $q(x,t)$ is proportional to the momentum. In quantum mechanics, momentum is not represented by differentiation with respect to time. It is obtained by differentiation of a spatial wave function with respect to spatial coordinates. Differentiation with respect to time yields energy.

The solution of eqn (8.20) will be restricted to waves that advance in the positive x direction. Thus, $e^{i(kx-\omega_k t)}$ and $e^{-i(kx-\omega_k t)}$ are used.[4] The fact that $e^{-i\omega_k t}$ and $e^{i\omega_k t}$ appear in these expressions does not result in a problem with negative kinetic energy, as plagued early attempts to construct a quantum mechanical theory consistent with special relativity. Recall that the Hamiltonian given by eqn (1.42) in Chapter 1: $\sum\omega_k(\hat{a}_k{}^\dagger\hat{a}_k+1/2)$ is always positive real. It has nothing to do with the oscillatory functions $e^{-i\omega_k t}$ and $e^{i\omega_k t}$.

So far, the equations that describe the time behavior are for the classical case. It is time to obtain the quantum version. To do this, the Heisenberg picture (Appendix 4) is used to assign time dependence to the operator $\hat{q}(x)$. In other words, $\hat{q}(x)$ becomes $\hat{q}(x,t)$ through the similarity transformation

$$\hat{q}(x,t) = e^{i\hat{H}t}\,\hat{q}(x)\,e^{-i\hat{H}t}, \tag{8.22}$$

and eqn (8.15) becomes

$$\hat{q}(x,t) = L^{-1/2}\sum_k \frac{1}{\sqrt{2\omega_k}}\left(e^{ikx}\left(e^{i\hat{H}t}\,\hat{a}_k\,e^{-i\hat{H}t}\right)+e^{-ikx}\left(e^{i\hat{H}t}\,\hat{a}_k{}^\dagger\,e^{-i\hat{H}t}\right)\right). \tag{8.23}$$

The right hand side of eqn (8.23) was obtained with the understanding that the Hamiltonian is given by eqn (8.10): $\sum\omega_k(\hat{a}_k{}^\dagger\hat{a}_k+1/2)$. This representation of the Hamiltonian depends on $\hat{a}_k$ and $\hat{a}_k{}^\dagger$, so it commutes with the terms e^{ikx} and e^{-ikx} in eqns (8.15) and (8.23). Thus, only $\hat{a}_k$ and $\hat{a}_k{}^\dagger$ are bracketed by $e^{i\hat{H}t}$ and $e^{-i\hat{H}t}$.

It is known that the Heisenberg picture, in which the operators evolve in time with the state vectors held fixed in time, comes closest to classical mechanics. The way the Heisenberg picture works is that a unitary transformation $e^{i\hat{H}t}$ is applied to each state vector. Because the time dependence of a Schrödinger state vector evolves according to $e^{-i\hat{H}t}$, the action of $e^{i\hat{H}t}$ removes the time dependence. This leaves a time independent state vector. In so doing, the operators acquire time dependence according to the prescription: $e^{i\hat{H}t}\hat{A}\,e^{-i\hat{H}t}$, where $\hat{A}$ is a time independent operator.

An intuitive way to see how this works is through consideration of the matrix element $\langle\,\psi_f(t)\,|\,\hat{A}\,|\,\psi_i(t)\,\rangle$. In Schrödinger quantum mechanics, $|\,\psi_i(t)\,\rangle$ evolves in time according to $|\,\psi_i(t)\,\rangle = e^{-i\hat{H}t}\,|\,i\,\rangle$, where $|\,i\,\rangle$ is time independent. Now assign $e^{-i\hat{H}t}$ to the operator to its left rather than the ket to its right. Doing the same with the $\langle\,\psi_f(t)|$ that lies to the left of $\hat{A}$ gives

---

[4] Were $kx+\omega t$ chosen, phase would advance in the opposite direction. That is, $d(kx+\omega t)/dt = 0$ gives a phase velocity $v_p = dx/dt = -\omega/k$. Wave packets travel in the $-x$ direction. On the other hand, $kx-\omega t$ has a phase velocity of ω/k and wave packets travel in the $+x$ direction.

$$\langle \psi_f(t) \,|\, \hat{A} \,|\, \psi_i(t) \rangle = \left(\langle f \,|\, e^{i\hat{H}t} \right) \hat{A} \left(e^{-i\hat{H}t} \,|\, i \rangle \right) = \langle f \,| \left(e^{i\hat{H}t} \hat{A}\, e^{-i\hat{H}t} \right) | \, i \rangle \,. \tag{8.24}$$

Thus, the time behavior is now associated with the operator, and the state kets are time independent. If you are unfamiliar with this approach, I suggest you look into it. It will not take long, and it is one of those things that appear from time to time. A good description of the Schrödinger, Heisenberg, and interaction pictures is in Sakurai's book: *Modern Quantum Mechanics* [41]. It is also given in Appendix 4 in reasonable detail, including applications. The algebra used to evaluate terms like $e^{i\hat{H}t}\hat{a}_k e^{-iHt}$ is straightforward and found in many books on quantum mechanics. In brief,

$$\begin{aligned} \hat{a}_k(t) \;&= e^{i\hat{H}t}\hat{a}_k e^{-iHt} = \left(1 + itH - t^2H^2 \ldots\right)\hat{a}_k\left(1 - itH - t^2H^2 \ldots\right) \\ &= \hat{a}_k + it\left[H, \hat{a}_k\right]\ldots \;= \hat{a}_k + it\omega_k\left[\hat{a}_k{}^\dagger \hat{a}_k, \hat{a}_k\right]\ldots \\ &= \hat{a}_k - it\omega_k \hat{a}_k \ldots \end{aligned} \tag{8.25}$$

The term $\hat{a}_k - it\omega_k\hat{a}_k \ldots$ is equal to $\hat{a}_k$ times the expansion of $e^{-i\omega_k t}$, though only the first two terms of the expansion are shown. Thus, the annihilation operator $\hat{a}_k$ has acquired time dependence according to

$$\boxed{\hat{a}_k(t) = \hat{a}_k e^{-i\omega_k t}\,.} \tag{8.26}$$

Likewise, $\hat{a}_k{}^\dagger(t)$ is given by $\hat{a}_k{}^\dagger e^{i\omega_k t}$. This is readily obtained by simply taking the adjoint of $\hat{a}_k(t)$, that is, $\hat{a}_k{}^\dagger(t) = (\hat{a}_k e^{-i\omega_k t})^\dagger = \hat{a}_k{}^\dagger e^{i\omega_k t}$. Note that even though the energy ω_k is positive we now have exponentials with positive and negative signs: $e^{-i\omega_k t}$ and $e^{i\omega_k t}$. However, $e^{i\omega_k t}$ comes about because an adjoint is taken, not because we admit a negative energy solution. As a result of the above maneuvers, we now have expressions for the field operators for displacement and momentum: $\hat{q}(x,t)$ and $\hat{p}(x,t)$

$$\hat{q}(x,t) = L^{-1/2} \sum_k \frac{1}{\sqrt{2\omega_k}} \left(e^{i(kx - \omega_k t)}\hat{a}_k + e^{-i(kx-\omega_k t)}\hat{a}_k{}^\dagger \right) \tag{8.27}$$

$$\hat{p}(x,t) = L^{-1/2} \sum_k -i\sqrt{\frac{\omega_k}{2}} \left(e^{i(kx - \omega_k t)}\hat{a}_k - e^{-i(kx-\omega_k t)}\hat{a}_k{}^\dagger \right). \tag{8.28}$$

8.4. Three-Dimensional Space

The extension of the results given by eqns (8.27) and (8.28) to 3D requires little work. Symbols, operations, and so on are simply replaced with their 3D counterparts: $x \to \vec{r}$; $k \to \vec{k}$; $\hat{q}(x,t) \to \hat{\phi}(\vec{r},t)$; $\partial^2 / \partial x^2 \to \nabla^2$; $L^{-1/2} \to L^{-3/2}$. In carrying this out, it is assumed that displacements in different directions are independent of one another. To get an idea of what this means, consider a stretched membrane, for example the surface of a drum. A point in the middle of the drumhead is displaced in the direction of the surface normal. The restoring force is aligned with the surface normal and directed oppositely to the displacement. However, the displacement also results in forces that are perpendicular to the surface normal. They happen to cancel one another when the point being displaced is at the center of a circular drumhead. Nonetheless, they exist and in general need to be taken into account.

Mathematically, this can be done using what is called the stress tensor. In our case these secondary forces are neglected. This, in effect, is equivalent to assuming small displacements, or the kind of medium for which displacements in different directions are truly independent. For the elastic limit, the small amplitude displacement assumption applies, while for other cases, displacements in different directions are independent.

Introducing the above substitutions yields the wave equation and the expression for the 3D field operator $\hat{\phi}(\vec{r},t)$:

$$\left(\frac{\partial^2}{\partial t^2} - v^2 \nabla^2 + \Omega_0^2 \right) \hat{\phi}(\vec{r},t) = 0 \tag{8.29}$$

and

$$\hat{\phi}(\vec{r},t) = L^{-3/2} \sum_k \frac{1}{\sqrt{2\omega_{\vec{k}}}} \left(e^{i(\vec{k}\cdot\vec{r} - \omega_{\vec{k}} t)} \hat{a}_k + e^{-i(\vec{k}\cdot\vec{r} - \omega_{\vec{k}} t)} \hat{a}_k{}^\dagger \right), \tag{8.30}$$

where $\omega_k{}^2 = \Omega_0{}^2 + v^2 k^2$, and $\omega_{\vec{k}}$ is taken to be positive.

Note that the displacement field operator given by eqn (8.30) is Hermitian: $\hat{\phi}(\vec{r},t) = \hat{\phi}^\dagger(\vec{r},t)$. Leaving aside the fact that $\hat{\phi}(\vec{r},t)$ is an operator, you are undoubtedly familiar with functions that have similar appearance, because they are used routinely to denote quantum mechanical wave functions. In the present context, however, $\hat{\phi}(\vec{r},t)$ is an *operator*, and a special one. It does not operate in coordinate or momentum space like the operators encountered in regular quantum mechanics. Instead, it operates in the number-valued space called Fock space. It creates and annihilates field quanta in single particle states given by the accompanying waves $e^{i(\vec{k}\cdot\vec{r} - \omega_{\vec{k}} t)}$ and $e^{-i(\vec{k}\cdot\vec{r} - \omega_{\vec{k}} t)}$. In this sense, it integrates the creation and annihilation of field quanta with regular quantum mechanics. The expression given by eqn (8.30) looks like a Fourier expansion in terms of waves, except that the expansion coefficients are operators.

8.5. Infinite Volume Limit

The infinite volume ($L \to \infty$) limit is obtained by replacing summation by integration. Summation is over the fundamental volume elements of $\vec{k}-$space : $(2\pi/L)^3$:

$$\sum_{\vec{k}} (2\pi/L)^3 \;\to\; \int d^3k \,. \tag{8.31}$$

In addition, the operators $\hat{a}_{\vec{k}}$ and $\hat{a}_{\vec{k}}{}^\dagger$ must be transformed to their continuum counterparts. This requires that we use: $\hat{a}(\vec{k}) = (2\pi/L)^{-3/2}\hat{a}_k$, and $\hat{a}(\vec{k})^\dagger = (2\pi/L)^{-3/2}\hat{a}_k{}^\dagger$. This is interesting because the factor $(2\pi/L)^{-3/2}$ diverges as $L \to \infty$. What is going on?

To see how the $(2\pi/L)^{-3/2}$ arises and ensure that nothing strange is going on with the math, consider the commutator $[\hat{a}(\vec{k}), \hat{a}(\vec{k}')^\dagger]$. The relation $\hat{a}(\vec{k}) = (2\pi/L)^{-3/2}\hat{a}_{\vec{k}}$ results in $[\hat{a}(\vec{k}), \hat{a}(\vec{k}')^\dagger]$ being equal to $(2\pi/L)^{-3}[\hat{a}_k, \hat{a}_{k'}{}^\dagger]$. However, we have seen earlier that $[\hat{a}_{\vec{k}}, \hat{a}_{\vec{k}'}{}^\dagger] = \delta_{\vec{k},\vec{k}'}$, and therefore

$$\left[\hat{a}(\vec{k}), \hat{a}(\vec{k}')^\dagger\right] = \left(\frac{L}{2\pi}\right)^3 \delta_{\vec{k},\vec{k}'} \,. \tag{8.32}$$

Now recall the discrete 1D lattice. The transformation to a continuous mass distribution resulted in the Kronecker delta $\delta_{n,n'}$ going over to the delta function $\delta(x-x')$ (see Fig. 4). That is, $\delta_{n,n'}/a$ in the limit: $a \to 0$ and $N \to \infty$, becomes $\delta(x-x')$. The same idea applies here. However, in $\vec{k}$-space the appropriate 3D volume element is $(2\pi/L)^3$. It is analogous to the a used in 1D coordinate space. Consequently, the limit that gives the delta function is

$$\lim_{L\to\infty} \frac{\delta_{k,k'}}{(2\pi/L)^3} = \delta(\vec{k}-\vec{k}') \,. \tag{8.33}$$

Therefore, the right hand side of eqn (8.32) is $\delta(\vec{k}-\vec{k}')$ and we have

$$\left[\hat{a}(\vec{k}), \hat{a}(\vec{k}')^\dagger\right] = \delta(\vec{k}-\vec{k}') \,. \tag{8.34}$$

When the above steps are taken into account, the field operator given by eqn (8.30) becomes

$$\hat{\phi}(\vec{r},t) = \int d^3k \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_k}} \left(e^{i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)}\hat{a}(\vec{k}) + e^{-i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)}\hat{a}(\vec{k})^\dagger \right). \tag{8.35}$$

As before, when this operator acts on a vector in Fock space it creates and annihilates field quanta. The integral form accommodates a continuous range of wave vectors.

In relativistic quantum field theory, it is necessary to accommodate the creation and annihilation of particles and their antiparticles. In the quantum mechanics we have dealt with in these notes, and that you have dealt with so far in your education, the number of massive particles is conserved. This is not the case in real life. For example, high-energy events can create pairs such as an electron and a positron, and do other interesting things. To achieve this in the mathematical formalism, it is necessary to have complex classical fields that become non-Hermitian quantum fields. The reason is that the field must have two components, one to account for a system's particle nature and the other to account for its antiparticle nature. In other words, the creation and annihilation terms must be independent of one another. Thus, $\hat{b}(\vec{k})^\dagger$ is substituted for $\hat{a}(\vec{k})^\dagger$:

$$\hat{\phi}(\vec{r},t) = \int d^3k \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_{\vec{k}}}} \left(e^{i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{a}(\vec{k}) + e^{-i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{b}(\vec{k})^\dagger \right).$$

This operator annihilates particles with wave vector $\vec{k}$ and creates antiparticles with wave vector $\vec{k}$. Its Hermitian adjoint does the reverse: creates particles with wave vector $\vec{k}$ and annihilates antiparticles with wave vector $\vec{k}$. This adjoint operator is

$$\hat{\phi}(\vec{r},t)^\dagger = \int d^3k \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_{\vec{k}}}} \left(e^{i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{b}(\vec{k}) + e^{-i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{a}(\vec{k})^\dagger \right).$$

We have come a long way from the model introduced at the beginning of Chapter 1.

Chapter 9.

Extension to Covariant Field Theory

Reeds
Robert Wittig

Contents

9.1. Introduction

The wide applicability of the results obtained using the model introduced at the start of Chapter 1 is amazing. The plasma served as a nice example. The quanta of the plasma field: bulk and surface plasmons, were obtained automatically. In the present chapter the model is extended to a different energy regime and physical system through straightforward manipulation. To begin, divide eqn (8.29) by v^2 to obtain

$$\left(\frac{\partial^2}{\partial (vt)^2} - \nabla^2 + \frac{{\Omega_0}^2}{v^2} \right) \hat{\phi}(\vec{r},t) = 0 \,. \tag{9.1}$$

Interesting substitutions are now made: $v \to c$ and $\Omega_0 \to mc^2$. Do not be alarmed by the appearance of discrete mass in mc^2. This energy term is a constant for a given particle type. The shorthand notation: $\partial_\mu \equiv \partial / \partial x^\mu$ is used, where $\mu = 0, 1, 2, 3$, with $x^0 = ct$, $x^1 = x$, $x^2 = y$, and $x^3 = z$. Equation (9.1) thus becomes

$$\left(\partial_0 \partial_0 - \nabla^2 + \left(\frac{mc}{\hbar} \right)^2 \right) \hat{\phi}(\vec{r},t) = 0 \,. \tag{9.2}$$

To be on the safe side, $\hbar$ is explicit in eqn (9.2). The term $\hbar / mc$ has units of length. For an electron its value is 0.00386 Å. It is referred to as the Compton radius (the Compton wavelength divided by 2π). The Compton wavelength is the wavelength of a photon that has the same energy as mc^2. In other words, $h\nu = hc/\lambda = mc^2 \implies \lambda = h/mc$. Therefore $\lambda / 2\pi = \hbar / mc$.

Equation (9.2) can be expressed compactly using a mathematical tool called the Minkowski metric, which handles relativistic spacetime economically, for example, with the squared interval: $(ct)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$. This quantity is a Lorentz scalar, and therefore observers in different inertial reference frames measure the same value for it. Its negative is also conserved and some people prefer to work with this. In any event, the fact that the squared interval returns the same value in any inertial frame leads to remarkable differences between the relativistic and non-relativistic regimes.

The Minkowski metric subsumes the essence of classical special relativity.[1] It will be introduced properly in Part V. For now, consider it a tool whose widespread use is a fact of life. In matrix form, it is

[1] The spacetime of special relativity is flat. This is not so for general relativity (gravity), where the metric assumes a more formidable form.

$$\eta^{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{9.3}$$

Repeated indices (one up, one down) imply summation – a convention introduced by Einstein. This contravariant form, $\eta^{\mu\nu}$, raises indices. Its covariant counterpart, $\eta_{\mu\nu}$, has the same $(+---)$ signature as $\eta^{\mu\nu}$ (same matrix representation), though it lowers indices.

The metric tensor facilitates moving freely between covariant and contravariant quantities. These appear naturally in special relativity. In fact, we have been using contravariant and covariant components all along and not recognizing it because there is no special significance in a Euclidean space, whereas, in general, there is great significance in relativistic spacetime. What I am going to do is borrow an explanation from Part V, and include it as an appendix to this chapter. You can skip it if you like, as we are essentially at the end of Part IV, but if you are interested it is there.

As mentioned above, the metric tensor $\eta_{\mu\nu}$ converts a contravariant four-vector to a covariant one. The simplest example is

$$x_\mu = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix} = \eta_{\mu\nu} x^\nu = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix}. \tag{9.4}$$

where $x^0 = x_0 = ct$.

From the above, we see that the Lorentz norm in this case is $x_\mu x^\mu = x^\mu x_\mu = (ct)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$. A clear description of the metric tensor is also given in Kusse and Westwig: *Mathematical Physics: Applied Mathematics for Scientists and Engineers* [36]. Thus, eqn (9.2), whose differential operator components are partial derivatives, is written

$$\left(\partial^\mu \partial_\mu + \left(\frac{mc}{\hbar} \right)^2 \right) \hat{\phi}(\vec{r},t) = 0 \,. \tag{9.5}$$

Notice that $\partial^\mu \partial_\mu$ is the d'Alembertian operator that was introduced earlier.

9.2. Klein-Gordon Equation

Equation (9.5) is referred to as the Klein-Gordon equation. It is a fundamental equation in relativistic quantum field theory. The development we have followed thus far has enabled us to obtain it in an expeditious manner. Most importantly, we have avoided the issues that plagued Oskar Klein (photo on next page) and Walter Gordon. They published independent papers in 1926-27 [70,71] that described their attempts to use eqn (9.5) as an equation in relativistic quantum mechanics. This is flawed fundamentally, but they pub-

lished their work anyway. It was disparaged roundly at the time, but it was resurrected in the mid-1930's when Pauli and Weisskopf showed that it is a correct equation in relativistic quantum field theory.

The reason for our good fortune is that we have approached eqn (9.5) from the very beginning as an equation in field theory. Moreover, we have enlisted a remarkably straightforward field theory: a balls and springs model. Having come this far, it should now be pointed out that a crucial feature was introduced at the outset with Fig. 1 of Chapter 1. Referencing the mass points to an inertial frame made it possible to eventually turn our balls and springs model into eqn (9.5). This enabled us to obtain first the plasma results and now the Klein-Gordon equation.

A number of interesting features can be seen with the result. For example, at fixed total energy, increasing the mass slows wave packets according to the dispersion relation indicated in Fig. 2. Such dependence is reasonable on mathematical grounds, though we usually do not think in these terms. Varying the mass is easy to do in a calculation, but with real systems mass is fixed, leaving aside the fact that at high energy, massive particles can be created and annihilated. We often use the fact that for a given energy the momentum of a heavy particle exceeds that of a light particle. For example, in the nonrelativistic regime, for given energy, linear momentum obeys: $p_h/p_l = (m_h/m_l)^{1/2}$, where h and l refer to heavy and light. Only infrequently, however, does one explore continuously variable particle mass.

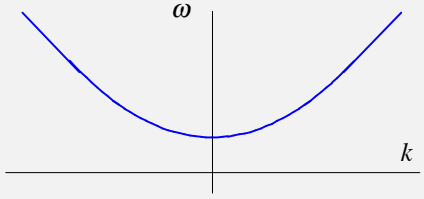


Figure 2. The dispersion relation: $\omega^2 = \Omega_0{}^2 + c^2k^2$ is hyperbolic. It has the same mathematical form (with changes $\Omega_0 \to mc^2$ and $v \to c$) as special relativity: $E^2 = m^2c^4 + c^2p^2$.

Special relativity gives $E^2 = m^2c^4 + c^2p^2$, which has the same mathematical form as the dispersion relation for our lattice model: $\omega^2 = \Omega_0{}^2 + c^2k^2$. Thus, the mass of an elementary particle can be seen as being equivalent to a fundamental frequency. Figure 2 shows that at small values of k the group velocity, $\partial\omega/\partial k$, is small, as the curve is rather flat. In this regime we can use

$$\omega = \Omega_0\left(1+\frac{v^2k^2}{\Omega_0{}^2}\right)^{1/2}. \tag{9.6}$$

The small-k region is obtained by retaining the first term in the expansion of the radical and introducing $\Omega_0 = mc^2$ and $v \to c$. This yields

$$\omega \approx \Omega_0 + \tfrac{1}{2}\frac{v^2k^2}{\Omega_0} \Rightarrow mc^2 + \tfrac{1}{2}\frac{c^2k^2}{mc^2}. \tag{9.7}$$

Using $p = k$ for a free particle, this becomes

$$\omega \approx mc^2 + p^2/2m \,. \tag{9.8}$$

We see that the kinetic energy of a particle in free space, $p^2/2m$, is recovered in the nonrelativistic limit. As energy increases, with Ω_0 (and therefore mass) fixed, the effect of the mass eventually becomes negligible because $\Omega_0 << ck$, and we have $\omega = ck$, which is none other than the de Broglie relation: $p = h/\lambda$. In this limit, the momentum is the same as that of a photon. Keep in mind that the field examined here is a scalar, and therefore its quantum has a spin of zero, whereas the photon is the quantum of a vector field, so it has spin of 1.

9.3. Non-Hermitian Fields

The expression for $\hat{\phi}(\vec{r},t)$ given in the last chapter as eqn (8.35) and reproduced below as eqn (9.9) is applicable to Hermitian fields.

$$\hat{\phi}(\vec{r},t) = \int d^3k \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_k}} \left(e^{i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{a}(\vec{k}) + e^{-i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{a}(\vec{k})^\dagger \right) \tag{9.9}$$

However, to accommodate particles and their antiparticles it is necessary to have complex classical fields that become non-Hermitian quantum fields. As mentioned in Chapter 8, the reason is that the field must have two degrees of freedom, one to account for a system's particle nature and the other to account for the system's antiparticle nature. In other words, the creation and annihilation terms in eqn (9.9) must be independent of one another in order to achieve the needed non-Hermitian nature.

Let us now rewrite eqn (9.9) with $\hat{b}(\vec{k})^\dagger$ substituted for $\hat{a}(\vec{k})^\dagger$:

$$\hat{\phi}(\vec{r},t) = \int d^3k \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_{\vec{k}}}} \left(e^{i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{a}(\vec{k}) + e^{-i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{b}(\vec{k})^\dagger \right) . \tag{9.10}$$

This field operator annihilates particles with wave vector $\vec{k}$ and creates antiparticles with wave vector $\vec{k}$. Its Hermitian adjoint does the reverse: creates particles with wave vector $\vec{k}$ and annihilates antiparticles with wave vector $\vec{k}$. This adjoint operator is given by

$$\hat{\phi}(\vec{r},t)^\dagger = \int d^3k \frac{1}{(2\pi)^{3/2}\sqrt{2\omega_{\vec{k}}}} \left(e^{i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{b}(\vec{k}) + e^{-i(\vec{k}\cdot\vec{r}-\omega_{\vec{k}}t)} \hat{a}(\vec{k})^\dagger \right) . \tag{9.11}$$

That is enough for now.

Appendix: Metric Tensor

This is taken from Part V. Wording has been modified in parts to fit the present chapter, but there have been no substantive changes. To begin, consider the 2D systems in Fig. 1: Entry (a) is the standard Cartesian frame. (b) In this skewed-axis frame, axis-parallel projections are used to obtain a^1 and a^2, which are referred to as the contravariant components. They are *defined* as axis-parallel projections. Their covariant basis vectors are labeled $\hat{g}_1$ and $\hat{g}_2$ to distinguish them from the basis vectors $\hat{e}_1$ and $\hat{e}_2$ ($\hat{e}_x$ and $\hat{e}_y$) of the Cartesian frame.

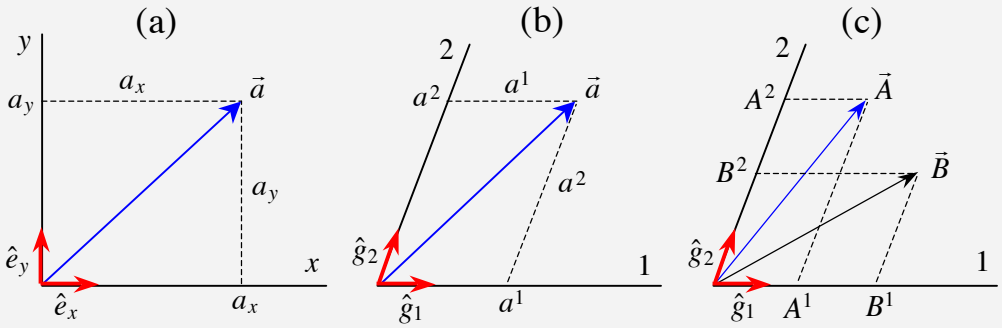


Figure 1. (a) Cartesian frame. (b) Skewed-axes frame showing axis-parallel projections. Superscripts and subscripts denote contravariant and covariant quantities, respectively. (c) Two vectors in the skewed-axes frame: Their inner product can be expressed in terms of contravariant components and covariant basis vectors [eqns (2) and (3)].

Next, consider the scalar product of two vectors $\vec{A}$ and $\vec{B}$ in a 2D space. In a frame with orthogonal basis vectors, this is written: $\vec{A}\cdot\vec{B} = A^i B^j \hat{e}_i \cdot \hat{e}_j$, with the basis vectors obeying $\hat{e}_i \cdot \hat{e}_j = \delta_{ij}$. If the axes are not orthogonal but skewed, as in Fig. 1(c), the same dot product operation yields

$$\left(A^1\hat{g}_1 + A^2\hat{g}_2\right)\cdot\left(B^1\hat{g}_1 + B^2\hat{g}_2\right) = A^1B^1\,\hat{g}_1\cdot\hat{g}_1 + A^2B^2\,\hat{g}_2\cdot\hat{g}_2 + A^1B^2\,\hat{g}_1\cdot\hat{g}_2 + A^2B^1\,\hat{g}_2\cdot\hat{g}_1 \tag{1}$$

If we define the $\hat{g}_i$ to be unit vectors, this becomes

$$\vec{A}\cdot\vec{B} = A^1\underbrace{\left(B^1 + B^2\,\hat{g}_1\cdot\hat{g}_2\right)}_{B_1} + A^2\underbrace{\left(B^2 + B^1\,\hat{g}_2\cdot\hat{g}_1\right)}_{B_2} \tag{2}$$

$$= A^1B_1 + A^2B_2\,. \tag{3}$$

The B_i in the above expression are referred to as the covariant components.

The geometric picture associated with the above equations is shown in Fig. 2. Let us focus on the B_1 component, as the B_2 component follows for like reason. In terms of contravariant components, it is written: $B_1 = B^1 + B^2 \hat{g}_1 \cdot \hat{g}_2 = B^1 + B^2 \cos\theta$. You can see from Fig. 2 and the above equation that the length of B_1 is given by the perpendicular projection of the vector $\vec{B}$ onto the 1-axis. Be careful, however. Figure 2 gives the lengths B_1 and B_2, but we do not yet have the associated basis vectors $\hat{g}^1$ and $\hat{g}^2$, and they certainly differ from $\hat{g}_1$ and $\hat{g}_2$. In other words, the quantities B_1 and B_2 are *not* aligned with the basis vectors $\hat{g}_1$ and $\hat{g}_2$, despite the fact that they are projections onto the 1 and 2 axes, respectively. This is subtle. Rather, they are aligned with $\hat{g}^1$ and $\hat{g}^2$, which are indicated in Fig. 3.

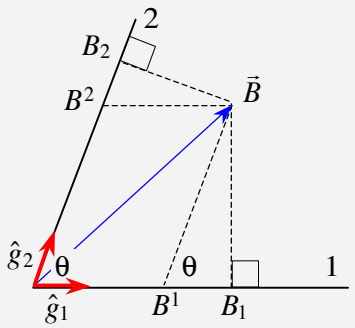


Figure 2. Covariant components are perpendicular projections onto the 1 and 2 axes.

In general, the covariant components can be obtained from the contravariant components through a straightforward operation. In the present example, this operation can be expressed in its most general form as

$$\begin{pmatrix} B_1 \\ B_2 \end{pmatrix} = \begin{pmatrix} \hat{g}_1 \cdot \hat{g}_1 & \hat{g}_1 \cdot \hat{g}_2 \\ \hat{g}_2 \cdot \hat{g}_1 & \hat{g}_2 \cdot \hat{g}_2 \end{pmatrix} \begin{pmatrix} B^1 \\ B^2 \end{pmatrix}. \tag{4}$$

It is understood that higher dimensional spaces follow suit. The 2×2 matrix in eqn (4), or its equivalent, is referred to as the metric. The metric need not be expressed as a matrix, but I thought this would be familiar to most of you. The metric for the orthogonal axis system in Fig. 1(a) is diag(1,1), so we need not pay attention to it, but for skewed axes it is important. To learn more about the covariant vector components, and particularly their contravariant basis vectors, we write $\vec{A} \cdot \vec{B}$ as

$$\left(A^1 \hat{g}_1 + A^2 \hat{g}_2\right) \cdot \left(B_1 \hat{g}^1 + B_2 \hat{g}^2\right) = A^1 B_1 \hat{g}_1 \cdot \hat{g}^1 + A^2 B_2 \hat{g}_2 \cdot \hat{g}^2 + A^1 B_2 \hat{g}_1 \cdot \hat{g}^2 + A^2 B_1 \hat{g}_2 \cdot \hat{g}^1 \tag{5}$$

Referring to eqn (3), we see that the right hand side of eqn (5) must be equal to $A^i B_i$. This requires that

$$\hat{g}_i \cdot \hat{g}^j = \delta_i{}^j. \tag{6}$$

We now have the geometric interpretation, which is indicated in Fig. 3. The covariant unit basis vectors $\hat{g}_1$ and $\hat{g}_2$ are orthogonal, respectively, to the (blue) contravariant basis vectors $\hat{g}^2$ and $\hat{g}^1$. The lengths of the contravariant basis vectors are established because their projections onto the covariant basis vectors, by definition, have unit length.

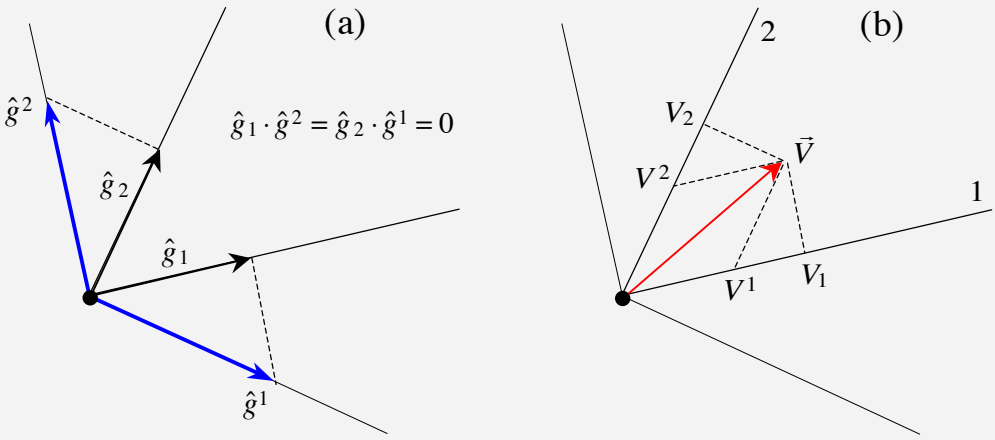


Figure 3. (a) The basis vectors obey the relation: $\hat{g}_i \cdot \hat{g}^j = \delta_i{}^j$, with the magnitudes of the covariant basis vectors $\hat{g}_i$ defined as unity. (b) The contravariant and covariant components are parallel and perpendicular projections, respectively.

The parallel-axis projections give the contravariant components V^1 and V^2, whereas the covariant components V_1 and V_2 are the perpendicular projections onto the axes. As mentioned earlier, the covariant components are taken from the axes of the covariant basis, rather than the axes of the contravariant basis. This can be confusing, but if you go through it a number of times it will sink in.

Let us now advance to special relativity and the Lorentz transformation. This time we shall go from unprimed Cartesian axes to primed skewed axes. The first odd thing we notice is that the unit basis vectors in the unprimed system appear to get bigger when they are transformed into the primed system, as seen for example with the transformation:

$$\begin{pmatrix} \hat{g}_{0'} \\ \hat{g}_{1'} \end{pmatrix} = \begin{pmatrix} \gamma & -\beta\gamma \\ -\beta\gamma & \gamma \end{pmatrix} \begin{pmatrix} \hat{g}_0 \\ \hat{g}_1 \end{pmatrix}. \tag{7}$$

When $\hat{g}_{0'} \cdot \hat{g}_{0'}$ is computed using $\hat{g}_{0'} = \gamma \hat{g}_0 - \beta\gamma \hat{g}_1$, we get

$$\hat{g}_{0'} \cdot \hat{g}_{0'} = \left(\gamma \hat{g}_0 - \beta\gamma \hat{g}_1\right) \cdot \left(\gamma \hat{g}_0 - \beta\gamma \hat{g}_1\right) \tag{8}$$

$$= \gamma^2 \hat{g}_0 \cdot \hat{g}_0 + \gamma^2 \beta^2 \hat{g}_1 \cdot \hat{g}_1 \tag{9}$$

$$= \gamma^2 \left(1 + \beta^2\right). \tag{10}$$

The value obtained, $\gamma^2(1+\beta^2)$, clearly exceeds unity; and likewise for $\hat{g}_{1'}$. The same occurs with other vector components such as *ct* and *x*.

The problem lies with the choice of metric. Computing $\hat{g}_{0'} \cdot \hat{g}_{0'}$ this way implies a metric of the form given by eqn (4), with standard (positive real) basis vectors. This works in many physical situations, but not here. In order to recover the squared interval of special relativity, it is necessary to use the Minkowski metric. The matrix representation is

$$\eta_{\mu\nu} = \begin{pmatrix} \hat{g}_0 \cdot \hat{g}_0 & 0 & 0 & 0 \\ 0 & \hat{g}_1 \cdot \hat{g}_1 & 0 & 0 \\ 0 & 0 & \hat{g}_2 \cdot \hat{g}_2 & 0 \\ 0 & 0 & 0 & \hat{g}_3 \cdot \hat{g}_3 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{11}$$

Everything falls into place. If we repeat the calculation of $\hat{g}_{0'} \cdot \hat{g}_{0'}$, but this time with the Minkowski metric, we get

$$\begin{bmatrix} \gamma, -\beta\gamma \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} \gamma \\ -\beta\gamma \end{bmatrix} = \gamma^2 \left(1 - \beta^2\right) = 1 . \tag{12}$$

Likewise, if we compute $\hat{g}_{1'} \cdot \hat{g}_{1'}$ using $\hat{g}_{1'} = -\beta\gamma \hat{g}_0 + \gamma \hat{g}_1$ we find

$$\begin{bmatrix} -\beta\gamma, \gamma \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix} \begin{bmatrix} -\beta\gamma \\ \gamma \end{bmatrix} = \gamma^2 \left(\beta^2 - 1\right) = -1 . \tag{13}$$

Other combinations do equally well. Let us now look further into the metric tensor.

The contravariant components Δx^0, Δx^1, Δx^2, and Δx^3 can be used to calculate the squared interval Δs^2. Once again, the math is subtle. For example, if we form the vector: $\Delta x^0 \vec{e}_0 + \Delta x^1 \vec{e}_1 + \Delta x^2 \vec{e}_2 + \Delta x^3 \vec{e}_3$ and take the usual Euclidean dot product of this vector with itself, a quantity is obtained that is not Lorentz invariant: $(\Delta x^0)^2 + (\Delta x^1)^2 + (\Delta x^2)^2 + (\Delta x^3)^2$. Clearly, this quantity does not match the Lorentz invariant squared interval Δs^2 or its negative, $-\Delta s^2$.

The modern way is to introduce imaginary basis vectors according to either of two options $\vec{e}_0$, $i\vec{e}_1$, $i\vec{e}_2$, $i\vec{e}_3$, or, alternatively, $i\vec{e}_0, \vec{e}_1, \vec{e}_2, \vec{e}_3$. We will use the former, so instead of $\vec{e}_1 \cdot \vec{e}_1 = 1$, we have $i\vec{e}_1 \cdot i\vec{e}_1 = -1$, and likewise $i\vec{e}_2 \cdot i\vec{e}_2 = -1$, and $i\vec{e}_3 \cdot i\vec{e}_3 = -1$. As mentioned earlier, an alternate (better) way of stating this is simply to define $\vec{e}_0 \cdot \vec{e}_0 = 1$, $\vec{e}_1 \cdot \vec{e}_1 = \vec{e}_2 \cdot \vec{e}_2 = \vec{e}_3 \cdot \vec{e}_3 = -1$, and $\vec{e}_\mu \cdot \vec{e}_\nu = 0$ for $\mu \neq \nu$.

This yields the metric tensor on Minkowski space: $\eta_{\mu\nu} = \vec{e}_\mu \cdot \vec{e}_\nu$, enabling Δs^2 to be calculated in a way that is analogous to the standard dot product on Euclidean space. To see how this works, consider the matrix representation of the metric tensor $\eta_{\mu\nu}$:

$$\eta_{\mu\nu} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix}. \tag{14}$$

When this acts on the contravariant components arranged in a column vector, the needed sign is obtained:

$$\begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} \Delta x^0 \\ \Delta x^1 \\ \Delta x^2 \\ \Delta x^3 \end{pmatrix} = \begin{pmatrix} \Delta x^0 \\ -\Delta x^1 \\ -\Delta x^2 \\ -\Delta x^3 \end{pmatrix} \tag{15}$$

$$= \begin{pmatrix} \Delta x_0 \\ \Delta x_1 \\ \Delta x_2 \\ \Delta x_3 \end{pmatrix}. \tag{16}$$

In this case, the required metric tensor is $\eta_{\mu\nu}$. It has two lower indices, making it the covariant representation. When it acts on the contravariant components Δx^ν, their indices are lowered and the covariant components Δx_ν are obtained. Thus, we see that $\Delta x_0 = \Delta x^0$, $\Delta x_1 = -\Delta x^1$, $\Delta x_2 = -\Delta x^2$, $\Delta x_3 = -\Delta x^3$. The entries in the column vectors on the right hand sides of eqns (15) and (16) are the covariant components.

Acting on the covariant components with $\eta^{\mu\nu}$ returns the column vector to its original contravariant representation. Thus, $\eta^{\mu\nu}$, which raises covariant components to contravariant ones, has the same form as eqn (14). The matrix representations of $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$ are the same, though it is understood that, respectively, they change contravariant to covariant and *vice versa*. The scalar Δs^2 is obtained by multiplying the right side of eqn (15) from the left by $[\,\Delta x^0\ \Delta x^1\ \Delta x^2\ \Delta x^3\,]$ to get $(\Delta x^0)^2 - (\Delta x^1)^2 - (\Delta x^2)^2 - (\Delta x^3)^2$.

A Few Manipulations Using $\eta_{\mu\nu}$ and $\eta^{\mu\nu}$

For the sake of brevity, nearly all of the Δ's are suppressed (taken as understood).

Lowering the indices of x^ν with the metric yields the covariant components:

$$\eta_{\mu\nu} x^\nu = x_\mu \Rightarrow \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix} = \begin{pmatrix} x_0 \\ x_1 \\ x_2 \\ x_3 \end{pmatrix}$$

Thus, the scalar Δs^2 is given by $x^\mu x_\mu = x^\mu \eta_{\mu\nu} x^\nu = (x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$. This is expressed in matrix form as:

$$\begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x^0 \\ x^1 \\ x^2 \\ x^3 \end{pmatrix} = \underbrace{\begin{pmatrix} x^0 & x^1 & x^2 & x^3 \end{pmatrix} \begin{pmatrix} x^0 \\ -x^1 \\ -x^2 \\ -x^3 \end{pmatrix}}_{(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2}$$

The same result is obtained by working from the other direction. Namely, use $x^\mu = \eta^{\mu\nu} x_\nu$ and form $\Delta s^2 = x_\mu x^\mu = x_\mu \eta^{\mu\nu} x_\nu = (x_0)^2 - (x_1)^2 - (x_2)^2 - (x_3)^2$. Note that this is equivalent to $(x^0)^2 - (x^1)^2 - (x^2)^2 - (x^3)^2$.

Part B. Bibliography and References

1. N. W. Ashcroft and M. D. Mermin, Solid State Physics (Brooks-Cole, Pacific Grove, CA, 1976).
2. J. S. Blackstone, Solid State Physics, Second Edition (Cambridge University Press, Cambridge, 1998).
3. J. R. Hook and H. E. Hall, Solid State Physics, Second Edition (Wiley, New York, 2004).
4. H. Ibach and H. Lüth, Solid State Physics (Springer Verlag, Berlin, 1993); this is an excellent all around reference.
5. C. Kittel, Introduction to Solid State Physics, 8th Ed. (Wiley, New York, 2005).
6. W. A. Harrison, Solid State Theory (Dover, New York, 1980).
7. M. A. Omar, Elementary Solid State Physics (Addison Wesley, New York, 1975).
8. O. Madelung, Introduction to Solid-State Theory (Springer, New York, 1996).
9. J. M. Ziman, Electrons and Phonons (Oxford University Press, Oxford, 2001).
10. J. M. Ziman, Theory of Solids (Cambridge University Press, Cambridge, UK, 1972).
11. M. Born and K. Huang, Dynamical Theory of Crystal Lattices (Clarendon Press, Oxford, 1954, reprinted 2002).
12. C. M. Wolfe, N. Holonyak, Jr., and G. E. Stillman, Physical Properties of Semiconductors (Prentice Hall, Englewood, New Jersey, 1989).
13. R. Hoffmann, Solids and Surfaces: A Chemist's View of Bonding and Extended Structures (Wiley VCH, New York, 1988).
14. L. Brillouin, Wave Propagation in Periodic Structures (Dover, Mineola, New York, 2003), reprinting of work published in 1946.
15. A. F. J. Levi, Applied Quantum Mechanics (Cambridge University Press, Cambridge, England, 2003).
16. R. Gilmore, Elementary Quantum Mechanics in One Dimension (Johns Hopkins University Press, Baltimore MD, 2004).
17. M. Fox, Optical Properties of Solids (Oxford University Press, Oxford, 2010); this is an excellent introductory level reference.
18. M. Fox, Quantum Optics: An Introduction (Oxford University Press, Oxford, 2010).
19. S. A. Maier, Plasmonics: Fundamentals and Applications (Springer, New York, 2007).
20. H. Raether, Surface Plasmons on Smooth and Rough Surfaces and on Gratings (Springer Tracts in Modern Physics 111, Berlin, 1988).
21. D. Sarid and W. Challener, Modern Introduction to Surface Plasmons (Cambridge University Press, Cambridge, 2010).

22. Shvets and Tsukerman, Plasmonics and Plasmonic Materials (World Scientific, Singapore, 2012).)
23. S. Kawata, Editor, Near-Field Optics and Surface Plasmon Polaritons (Springer Topics in Applied Physics, Volume 81, New York, 2010).
24. R. Claus, L. Merten, and J. Brandmüller, Light Scattering by Phonon Polaritons (Springer-Verlag Tracts in Modern Physics 75, Berlin, 1975).
25. I. D. Mayergoyz, Plasmon Resonances in Nanoparticles (World Scientific, Singapore, 2013).
26. M. Pelton and G. Bryant, Introduction to Metal Particle Plasmonics (John Wiley and Sons, Hoboken, New Jersey, 2013)
27. E. L. Albuquerque and M. G. Cottam, Polaritons in Periodic and Quasiperiodic Structures (Elsevier B. V., Amsterdam, 2004)
28. D. J. Amit and Y. Verbin, Statistical Physics: An Introductory Course (World Scientific, Singapore, 1999).
29. P. Coleman, Introduction to Many Body Physics, e-version available now, to be published October 2015.
30. J. R. Schrieffer, Theory of Superconductivity (Persus Books, Reading, MA, 1983).
31. P. Nozières, translation by D. Hone, Theory of Interacting Fermi Systems (W. A. Benjamin, New York, 1964).
32. A. Pines and P. Nozières and D. Pines, The Theory of Quantum Liquids (Perseus Books, Cambridge MA, 1999).
33. J. D. Jackson, Classical Electrodynamics, 3rd Edition (Wiley, New York, 1999).
34. E. M. Purcell, Electricity and Magnetism, 2nd Edition (McGraw-Hill, New York 1985).
35. G. B. Arfken and H. J. Weber, Mathematical Methods for Physicists, 6th Edition (Academic Press, New York, 2012). There are many editions of this book, though they hardly differ in content.
36. B. Kusse and E. Westwig, Mathematical Physics: Applied Mathematics for Scientists and Engineers (Wiley, New York, 2010). This is an excellent all-around reference. It explains nicely the geometry of covariant and contravariant components and basis vectors together with associated transformations.
37. J. Matthews and R. L. Walker, Mathematical Methods of Physics, Second Edition (Addison-Wesley, New York, 1970).
38. M. J. Lighthill, Introduction to Fourier Analysis and Generalized Functions (Cambridge University Press, Cambridge, 1958).
39. J. W. Brown and R. V. Churchill, Complex Variables and Applications, Ninth Edition (McGraw-Hill, New York, 2013).
40. G. Baym, Lectures on Quantum Mechanics (Addison Wesley, New York, 1990).
41. J. J. Sakurai, Modern Quantum Mechanics (Addison-Wesley, New York, 1994).
42. J. J. Sakurai, Advanced Quantum Mechanics (Addison-Wesley, New York, 1964).

43. F. Schwabl, Advanced Quantum Mechanics (Springer-Verlag, Berlin, 2006).
44. W. L. Mochán, Centro de Ciencias Físicas, Universidad Nacional Autónoma de México, Apartado Postal 48-3, Cuernavaca, Mor., 62251 México. Downloaded from web.
45. E. Kretschmann and H. Raether, Z. Naturforschung **23A**, 2135 (1968).
46. A. Otto, Z. Physik **216**, 398 (1968).
47. J. Lindhard, K. Dan. Vidensk. Selsk. Mat. Fys. Medd. **28**, 8 (1954).
48. P. Debye and E. Hückel, Physikalische Zeitschrift **24**, 185-206 (1923).
49. H. Amano, 2014 Physics Nobel Prize Lecture: Growth of GaN on Sapphire by Low Temperature Deposited Buffer Layer and Realization of p-type GaN by Mg-doping Followed by LEEBI Treatment.
50. W. G. Spitzer and H. Y. Fan, Phys. Rev. **106**, 882 (1957).
51. R. D. Averitt, Nature Physics **4**, 14 (2008).
52. C. Jagannath, Z. W. Grabowski, and A. K. Ramdas, Linewidths of the Electronic Excitation Spectra of Donors in Silicon, Phys. Rev. B **23**, 2023-2098 (1981).
53. G. Mie, Annalen der Physik **330**, 377-445 (1908). (1908)
54. Chad A. Mirkin is a world-renowned nanoscience expert, known for his development of nanoparticle-based biodetection schemes, the invention of Dip-Pen Nanolithography, and contributions to supramolecular chemistry.
55. Younan Xia, from Nanotechweb.org interview, December 2002. At the time, Xia was at the University of Washington; he has since moved to Georgia Tech, where he runs one of the most productive research groups in the world in this area.
56. R. Fuchs, Theory of the Optical Properties of Ionic Crystal Cubes, Phys. Rev. B **11**, 1732-1740 (1975).
57. E. Ringe, J. M. McMahon, K. Sohn, C. Cobley, Y. Xia, J. Huang, G. C. Schatz, L. D. Marks, and R. P. VanDuyne, J. Phys. Chem. C **114**, 12511-12516 (2010).
58. M. S. Dresselhaus, Solid State Physics Part II, Optical Properties of Solids, download from web.
59. H. Raether, Excitation of Plasmons and Interband Transitions by Electrons (Springer-Verlag, Berlin, 1980).
60. V. M. Agranovich and D. L. Mills, Surface Polaritons: Electromagnetic Waves at Surfaces and Interfaces (North Holland, New York, 1982).
61. R. H. Bube, Electrons in Solids: An Introductory Survey, Third Edition (Academic Press, New York, 1992).
62. A. Verma, Dayalbagh Educational Institute, Agra, Uttar Pradesh, India.
63. H. Goldstein, C. Poole, and J. Safko, Classical Mechanics, Third Edition (Pearson, New York, 2002).
64. J. B. Marion and S. T. Thornton, Classical Dynamics, Fourth Edition (Harcourt, New York, 1995).

65. P. Lorrain, D. R. Corson, and F. Lorrain, Electromagnetic Fields and Waves, Third Edition (W. H. Freeman and Company, New York, 1988).

66. R. Lyddane, R. Sachs, and E. Teller, On the Polar Vibrations of Alkali Halides, Phys. Rev. **59**, 673-676 (1951).

67. P. Andrade and S. P. S. Porto, Revista Brasileira de Física **3**, 337-343 (1973).

68. P. A. M. Dirac, Proc. Roy. Soc. (London) **A117**, 610 (1928); ibid. **A118**, 351 (1928).

69. M. J. Lighthill, Introduction to Fourier Analysis and Generalized Functions (Cambridge University Press, Cambridge UK, 1958).

70. O. Klein, Z. Phys. **37**, 895-906 (1926).

71. W. Gordon, Z. Phys. **40**, 117-133 (1926-1927).

72. I. J. R. Aitchison and A. J. G. Hey, Gauge Theories in Particle Physics, Volume I: From Relativistic Quantum Mechanics to QED (Taylor and Francis, New York, 2004).

73. I. J. R. Aitchison and A. J. G. Hey, Gauge Theories in Particle Physics, Volume II: QCD and the Electroweak Theory (Taylor and Francis, New York, 2004).

74. M. Guidry, Gauge Field Theories: An Introduction with Applications (Wiley Interscience, New York, 1999).

75. T. Baer and W. L. Hase, Unimolecular Reaction Dynamics (Oxford University Press, Oxford UK, 1996).

Part B. Exercises and Projects

Most of the exercises in Part B are embedded in the text. These are not repeated here. Those listed below are additional ones that accrued over the past half dozen years, plus a few projects that were assigned as presentations.

Chapter 1

1. Given $H = p^2/2m + kx^2/2 - ax$ (where a is small and positive): (i) sketch the potential showing how the $-ax$ term affects the shape near the bottom of the well; (ii) express the H matrix for the 5 lowest harmonic oscillator basis functions.

2. Consider the $|v\rangle$ harmonic oscillator basis. Which, if any, of the following matrix elements are zero: $\langle v|\hat{x}^2|v+2\rangle$; $\langle v|\hat{x}\hat{p}|v+2\rangle$; $\langle v|\hat{x}^2\hat{p}^2|v+2\rangle$?

3. A particle in a harmonic potential is in a state ψ that is not an eigenstate of the Hamiltonian. Instead, $\psi = 3^{-1/2}(\psi_0 + \psi_1 + \psi_2)$, where ψ_0, ψ_1, and ψ_2 are the three lowest energy harmonic oscillator eigenfunctions. Obtain the expectation value of $\hat{x}^2$: $\langle\psi|\hat{x}^2|\psi\rangle$.

4. We want the uncertainty product $\Delta x \Delta p$ for any eigenstate of the harmonic oscillator. To achieve this, evaluate $\langle v|\hat{x}^2|v\rangle$ and $\langle v|\hat{p}^2|v\rangle$ and use these results to evaluate $\Delta x \Delta p$.

5. Here we shall use the fact that the dispersion relation is not quantum mechanical. The dispersion relation is given by the eigenvalues obtained using the classical expression in eqn (1.1): $m\ddot{q}_n = -\kappa_0 q_n - \kappa(2q_n - q_{n-1} - q_{n+1})$. For sinusoidal oscillation the left hand side becomes $-m\omega^2 q_n$. To simplify matters assume $\kappa_0 = 0$. Set up a matrix representation of this expression using the displacements as basis vectors and diagonalize it to obtain the frequencies. Make a plot of frequency versus wave vector. Then consider $\kappa_0 \neq 0$.

6. Consider a linear chain of identical masses m, each connected to its nearest neighbors with springs whose constant is κ. There is a damping force: $-m\gamma\dot{q}_n$, where q_n is the displacement of the n^{th} particle from its equilibrium position. How does this damping change the frequencies? What is the relaxation rate of the modes (displacements)? Assume $\gamma^2 << \kappa/m$. Discuss the two extremes: $k \approx \pi/a$, and $k \approx 0$.

7. The equations of motion that go along with Fig. 1 are given by eqn (1.1). Go through the steps leading to eqn (1.2).

8. Expansions for p_n and P_k are given in eqns (1.33) and (1.34). Given p_n, show how inverting this yields the expression for P_k. Why does p_n have a negative exponent, whereas q_n has a positive exponent?

9. Referring to eqn (1.40), derive all of the commutator relations for the creation and annihilation operators.

10. This problem illustrates how the mathematics works for the commutator $[\hat{Q}_k, \hat{P}_{k'}]$. Take the case of just 3 sites: $n = 1, 2, 3$. This is not at all like the large-N loop in Fig. 3, but it will make the point. The normal coordinates and momenta are

$$\hat{Q}_k = \left(\frac{m}{3}\right)^{1/2}\left(q_1 e^{-ika} + q_2 e^{-i2ka} + q_3 e^{-i3ka}\right)$$

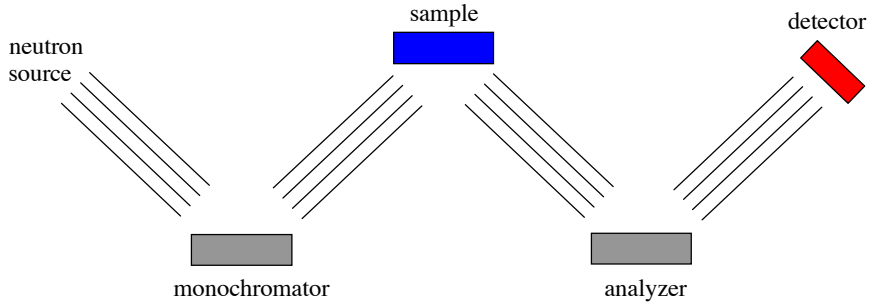
$$\hat{P}_{k'} = \left(\frac{1}{3m}\right)^{1/2}\left(p_1 e^{ik'a} + p_2 e^{i2k'a} + p_3 e^{i3k'a}\right)$$

Express $[\hat{Q}_k, \hat{P}_{k'}]$ as the sum of 9 commutators. Of these, which ones vanish? Explain; show how the answer is obtained. After you have done $N = 3$, the case $N = 2$ should be obvious. But what about $N = 1$? Show how this works. Discuss the issue of Hermiticity for $N = 1$ and $N > 1$.

11. In most quantum mechanics textbooks, the harmonic oscillator is solved through the introduction of raising and lowering (creation and annihilation) operators. For example, beginning with $\hat{H} = \hat{p}^2/2m + \kappa\hat{x}^2/2$ and the definition of the raising and lowering operators, this Hamiltonian assumes the form: $\hat{H} = \omega_0(\hat{a}^\dagger\hat{a} + 1/2)$, where $\omega_0{}^2 = \kappa/m$. In this expression, the operator $\hat{a}^\dagger\hat{a}$ is similar to the analogous expression developed in this chapter, $\hat{a}_k{}^\dagger\,\hat{a}_k$. Write a brief description of the relationship between the two approaches and results.

Chapter 2

1. A triple axis spectrometer consists of three independently controlled axes of rotation for the sample, monochromator, and analyzer crystals, as shown below. Inelastic neutron scattering is carried out with a beam of neutrons whose wavelength is 2.5 ± 0.002 Å. The analyzer crystal provides wavelength selection with accuracy $\delta\lambda/\lambda = 10^{-3}$. What is the energy resolution (in eV) of the spectrometer? Repeat for x-rays of the same wavelength.



2. What is the speed of sound (acoustic branch) for $\Omega = 10\ \text{cm}^{-1}$ and $a = 5$ Å?

3. Equation (2.22) was derived using the low-temperature limit of a dispersion relation. Later the Debye model yielded an expression valid in the low temperature limit. How do these results compare? Specifically, if they differ, explain why, and if they are the same, explain why.

4. Calculate $\langle E \rangle$ and C for the following quantum mechanical systems, each at temperature T: (a) two harmonic oscillators (frequencies ω_1 and ω_2); and (b) a system with just two energy levels. Discuss the difference between these cases. Discuss high and low temperature limits. Give an example of (b).

5. Use the Debye model to calculate C for a 2D monatomic crystal. This requires that you determine k_D that contains the same number of states as the Brillouin zone.

6. Two monatomic crystals (labeled 1 and 2) are each 1 cm × 1 cm × 1 cm. They have identical masses. However, their respective speeds of sound (averaged over the three acoustic branches) are $v_1 = 1000$ and $v_2 = 2000$ m/s. The temperatures of crystal 1 and crystal 2 are set to 10 and 20 K, respectively, and then the crystals are brought together in high vacuum such that they share a common side. What is the temperature of the combination after equilibrium has been reached? Explain your reasoning.

7. A graduate student needs to measure the temperature gradient along a sapphire rod that is connected at one end to a high temperature and at the other end to a low temperature. The student carries out the needed measurements and discovers that temperature falls off exponentially along the length. She brings this result to her advisor. They publish the result in *Science*, with a lot of fanfare. What do you think: does this result make sense; might it be important; might it be wrong? Hint: Use the divergence theorem to assess the flux along the length.

8. Referring to eqns (2.38) and (2.39), it was shown that $P = 0$ for a phonon. Show that the expectation value of the k-space operator $\hat{P}_k$ vanishes: $\langle n_k | \hat{P}_k | n_k \rangle = 0$. Now show that $\langle n_k | \hat{P}_k \hat{P}_k^{\dagger} | n_k \rangle \neq 0$. What is its value?

9. The equation of motion for an anharmonic oscillator is: $m\ddot{u} = -\kappa u + g u^2$. Solve for u by using the fact that it can be expanded in a harmonic series:

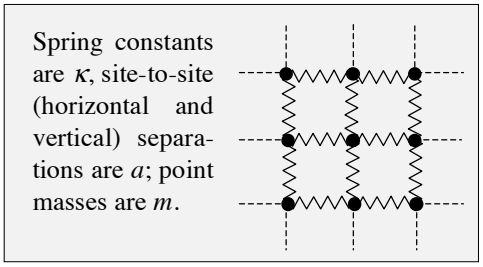
$$u = \sum_{n=1}^{\infty} c_n e^{in\omega_0 t} \ .$$

10. The Debye model uses a crude approximation for the dispersion relation. Improvement is achieved with the dispersion relation:

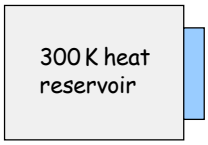
$$\omega = vk\left(1 - \frac{ka}{2\pi}\right),$$

where v is the speed of sound and k is positive. (a) Show that the group velocity vanishes at the edge of the Brillouin zone. (b) Using the above equation, obtain $k(\omega)$. (c) Calculate the 1D k-space density. Plot it. (d) Calculate the 2D k-space density.

11. Derive the k-space density, $\rho_2(k)$, where k is the magnitude of $\vec{k}$, for the 2D monatomic lattice indicated in the box. Discuss the couplings that exist between displacements and motions in the horizontal and vertical (x_1 and x_2) directions. Write the equation of motion for either direction. Be careful with the approximations that need to be made.



12. A sequence of experiments is carried out in which three different thermally conducting circular slabs (light blue in the figure on the right) are connected (one at a time) to a 300 K heat reservoir as shown. The slabs all have the same shape: 10 cm thickness and a radius that is much larger than the thickness. At t = 0 the right hand side of a given slab is connected to a 0 K heat sink that remains at this temperature thereafter. How long does it take for the temperature to reach 150 K in the middle of the slab (5 cm from either side) for each of the three materials used in the experiments: copper, gold, and diamond? What about orientation?



Chapter 4

1. Work out the expression for the plasma frequency for the case of the plasma slab shown in Fig. 9.

2. An intrinsic semiconductor slab is irradiated uniformly with the output of a pulsed laser (1 ps). This creates an effectively instantaneous electron density in the conduction band. The t = 0 value of the conduction band electron density is $n_e = 10^{21}$ cm^{-3}. The plasma thus created decays through spontaneous emission, i.e., an electron goes from the conduction band to the valence band and a photon is emitted. The radiative (spontaneous emission) lifetime is 10 ns. Assume a lattice dielectric constant of $\varepsilon_{lat} = 9$, and that the effective mass of the conduction band electrons is that of a bare electron. Ignore electron momentum relaxation, any contributions from holes, and any contributions from lattice phonons.

(a) What is the spectrum of the emitted radiation?

(b) An infrared laser beam is now passed through the slab and detected on the other side. It is a CW (continuous wave) laser, and its wavelength is 10 μm. Plot the detected intensity versus time starting at $t = 0$.

Chapters 6 and 7

1. Can an LO phonon radiate photons? If no, why not? If yes, in what direction do the photons travel relative to the LO phonon wave vector?

2. Can polaritons exist near the $\pm\pi/a$ edges of the Brillouin zone? If so, what photon energy is required?

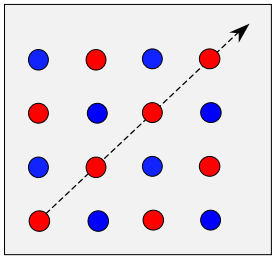
3. The phonon polariton curves look like an avoided crossing in k-space. What are the curves that cross? What is the interaction term that couples the curves?

4. When an electromagnetic wave passes through a lossless, bulk plasma can it be absorbed by the plasma, for example, if its frequency is close to the plasma frequency?

5. Consider the LO and TO modes of a NaCl crystal near $k = 0$. Which of these modes is closest in character to the vibration of the gas phase NaCl diatom. What considerations motivated your choice.

6. Electrons, being fermions, enter states one at a time. For example, the highest energy electrons in the conduction band of a metal typically have energies of several eV. Only those near the Fermi level participate in thermal conductivity. The rest are said to be "frozen out." Now consider a plasma made of conduction band electrons. Do all of these electrons participate in plasma oscillation? Does a subset participate; if so, which subset? Explain.

8. A 3D cubic ionic crystal is comprised of red and blue atoms as indicated with the sketch on the right of a cross section of the crystal. An experimentalist causes the atoms to vibrate. Do not worry about how she achieves this. Just take for granted that it happens. She then measures photons that appear at the angle relative to the crystal reference frame indicated by the arrow. Explain what is happening in the crystal in terms of its LO and TO modes.



Chapter 8

1. Equation (8.9) gives an expression for the amplitude $\hat{q}(x)$. Obtain the corresponding expression for $\hat{p}(x)$ by using similar manipulations.

2. Solve eqn (8.20) to obtain the space and time dependencies of the classical fields.

3. Consider eqns (8.27) and (8.28). Calculate the commutator of the field operator with its time derivative: $[\hat{\phi}(\vec{r},t), \dot{\hat{\phi}}(\vec{r}\,',t)]$. What do you make of this result? That is, how is one to interpret it? What about $[\hat{\phi}(\vec{r},t), \dot{\hat{\phi}}(\vec{r},t')]$? What about $[\hat{\phi}(\vec{r},t), \dot{\hat{\phi}}(\vec{r}\,',t')]$?

Projects

1. *Diatomic lattice*: Review the ionic lattice. There, the ion masses differed ($m_A \neq m_B$) and all inter-ion distances were the same, a. Now write a model for a 1D diatomic lattice in which the masses are the same, but there are two harmonic constants, κ_0 and κ. For example this might apply to a chain of oxygen molecules. Obtain the dispersion relation and make a figure showing displacements for the acoustic and optical branches in the region $k \sim 0$ and the zone edge. Show what happens for a broad range of force constants, that is, at one extreme values of κ_0 and κ differ greatly from one another; at the other extreme they each have the same value.

2. *Infrared spectrum of a molecular crystal*: The dispersion relation for the 1D lattice is examined, and an infrared spectrum is calculated. Referring to Fig. 8 in Chapter 1, choose the frequencies to be $\Omega_0 = 2300$ cm^{-1} and $\Omega = 10$ cm^{-1}. These values, were no adjustment made to the vertical scale, would result in a figure that appears quite different than Fig. 8. Make a plot using the values $\Omega_0 = 2300$ cm^{-1} and $\Omega = 10$ cm^{-1}. Assume $a = 5$ Å. Choose the number of sites to be large – your choice. Be sure to add a break to the vertical scale to ensure that the plot is readable.

   The value 2300 cm^{-1} is similar to the CO_2 asymmetric stretch frequency (ν_3), and the value 10 cm^{-1} is similar to the vibrational frequency of a van der Waals bond. Incidentally, crystalline CO_2 does not resemble a linear chain of CO_2 molecules laid end-to-end: $OCO\cdot\cdot OCO\cdot\cdot OCO\cdot\cdot$. However, we are dealing with a model whose purpose is to reveal principles and trends so do not worry about the real structure.

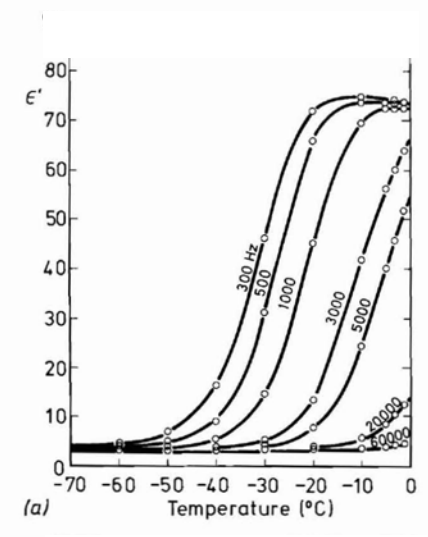
   Now calculate the spectrum for a 100 K sample. Temperature enters only at the end when a thermal distribution of bosons is introduced. The transition moment is dominated by the ν_3 mode. This is a strong infrared absorption (CO_2 is *the* greenhouse gas). What is the selection rule for k? In other words, what is the relationship between the k value of the lower level and the k value of the upper level? Explain how this works. Plot the spectrum as intensity versus frequency (this is where temperature comes in). Finally, use a range of parameter values in order to see how the system varies with temperature and with the van der Waals frequency, Ω.

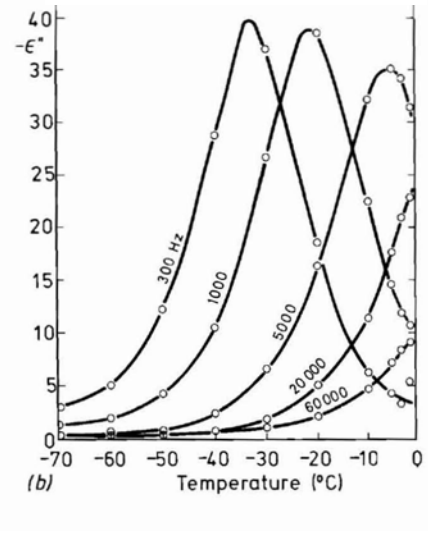
3. *Surface enhancement*: Surface plasmons form polaritons that have unique properties. One aspect of plasmons is their high polarizability. In this project, you are to present a critical summary of surface enhanced optical processes on metal surfaces, including molecules (up to monolayers) adsorbed on metal surfaces.

4. *Polaritons*: Start by examining the phonon polariton. To assess how the strength of the interaction between the electromagnetic field and the crystal is manifest in the properties of the system, plot the dispersion relation for $\varepsilon_\infty = 4$ and ε_0 values of 5,

6, 8, and 10. Discuss the result. Now calculate the respective fractions of mechanical (that is, vibrational) and electromagnetic energies as ω varies from below to above the ω_T resonance. Explain what a phonon polariton means to you in intuitive, physical terms, namely, without resorting to equations. Discuss a surface plasmon polariton. Plot the dispersion relation and give a qualitative description of what is going on.

5. *Low frequency modes of ice*: Consider the dispersion relation for an ionic lattice. Add a loss term to the equation of motion and derive the expression for dielectric function versus frequency. Make plots that show how the real and imaginary parts vary. Discuss the significance.





Now consider crystalline ice. The ionic model does not take into account the low frequency modes of a molecular solid. Namely, the macroscopic polarization can be diminished by motion of the H_2O units that move the induced dipole around in space. This can be represented with a phenomenological equation:

$$\frac{d\boldsymbol{P}}{dt} = -\frac{\boldsymbol{P}}{\tau} + \frac{1}{\tau}\alpha(0)\boldsymbol{E}_0 e^{-i\omega t}$$

where $\alpha(0)$ is a quasi-static polarizability, and τ is the phenomenological relaxation time. Assume there is no net macroscopic polarization in the absence of the applied field. Use the above equation with $\boldsymbol{P} = \alpha \boldsymbol{E}$ to obtain a dispersion relation. Hint: Fourier transformation of the above equation gives a complex polarization. Make a plot of the real and imaginary parts of the dielectric function. Now consider the data in the figure. Fit these to a relaxation time that varies with temperature according to:

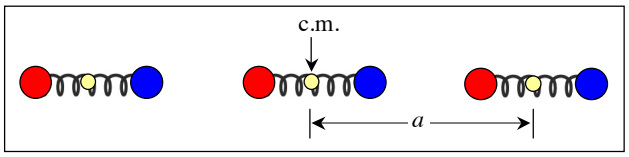
$$\tau = \tau_0 e^{T_0/T}$$

What is the significance of this fit? On the basis of this model, what prediction can you make for temperatures in the range 150-200 K? Can this information be used to understand the morphological changes that take place in amorphous solid water?

6. *Wave packets*: Consider the dispersion relation for a 1D ionic lattice consisting of 50 diatom pairs. Construct a localized $t = 0$ wave packet that is appropriate to the acoustic branch. Use a range of k values that obeys (approximately) $\omega_k = vk$. In other words, stay away from the edge of the first Brillouin zone. Make a figure showing $|\psi(t)|^2$. Use realistic values for parameters. Show how this packet propagates by

calculating $|\psi(t)|^2$ at different times (snapshots). Repeat the calculations for the optical branch. Compare the low-k, mid-k, and zone edge regions. Discuss the results.

7. *Ferroelectric phase transition*: Consider a 1D lattice of diatomic molecules separated by a distance a. Assume their centers-of-mass (yellow dots) remain fixed in space, and the diatom frequency is ω_0. The molecules have permanent dipole moments and they are polarizable, with polarizability α. The dipole moment of the diatom at site m is expressed as $p_m = \bar{p}e^{i(kam-\omega t)}$, where $\bar{p}$ is a constant. The electric field at site n due to the dipole at site m has a magnitude $2p_m|a(n-m)|^{-3}$, and its direction is parallel to the axis of the lattice. This field contributes to the dipole moment at site n through the polarizability.



Calculate the field at site n due to all of the diatomic molecules (sum the contributions). The equation of motion for the dipole at site m is: $\ddot{p}_m = -\omega_0^2 p_m + \alpha E \omega_0^2$, where E is the local electric field. Explain the terms. Obtain the dispersion relation. Find the critical value of the polarizability, α_{crit}, beyond which k becomes imaginary. Discuss this ferroelectric phase transition.

Part IV. Appendices